

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Boynton, John E.
Gillham, Nicholas W.
Randolph-Anderson, Barbara L.
Ishige, Fumiharu
Sato, Ryo
- (ii) TITLE OF INVENTION: Methods of Conferring PPO-Inhibiting
Herbicide Resistance to Plants by Gene Manipulation
- (iii) NUMBER OF SEQUENCES: 24
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
(B) STREET: P.O. Box 747
(C) CITY: Falls Church
(D) STATE: Virginia
(E) COUNTRY: USA
(F) ZIP: 22040-3487
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: Patent In Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US new
(B) FILING DATE: 30-SEP-1996
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Murphy Jr., Gerald M.
(B) REGISTRATION NUMBER: 28,977
(C) REFERENCE/DOCKET NUMBER: 2185-156P
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 703-205-8000
(B) TELEFAX: 703-205-8050

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO

00347 081660

65

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Chlamydomonas reinhardtii*

(B) STRAIN: CC-407

(ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION: 1..47

(D) OTHER INFORMATION: /product= "porphyric herbicide resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ala	Ala	Glu	Ala	Leu	Gly	Ser	Phe	Asp	Tyr	Pro	Pro	Val	Gly	Ala	Val
1				5					10					15	

Thr	Leu	Ser	Tyr	Pro	Leu	Ser	Ala	Val	Arg	Glu	Glu	Arg	Lys	Ala	Ser
			20					25					30		

Asp	Gly	Ser	Val	Pro	Gly	Phe	Gly	Gln	Leu	His	Pro	Arg	Thr	Gln
		35					40					45		

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(B) STRAIN: ecotype Columbia

(ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION: 1..46

(D) OTHER INFORMATION: /product= "porphyric herbicide resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ala	Ala	Asn	Ala	Leu	Ser	Lys	Leu	Tyr	Tyr	Pro	Pro	Val	Ala	Ala	Val
1				5					10					15	

Ser	Ile	Ser	Tyr	Pro	Lys	Glu	Ala	Ile	Arg	Thr	Glu	Cys	Leu	Ile	Asp
			20					25					30		

003312272260

66

Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Thr Gln
 35 40 45

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Zea mays*
- (B) STRAIN: B73 inbred

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..46
- (D) OTHER INFORMATION: /product= "porphyric herbicide resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ala Ala Asp Ala Leu Ser Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val
 1 5 10 15
 Thr Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp
 20 25 30
 Gly Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln
 35 40 45

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 141 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Chlamydomonas reinhardtii*
- (B) STRAIN: CC-407

09331723-081899

67

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..141
- (D) OTHER INFORMATION: /note= "encodes porphyric herbicide resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
GCCGCCGAGG CCCTGGGCTC CTTGACTAC CCGCCGGTGG GCGCCGTGAC GCTGTCGTAC      60
CCGCTGAGCG CCGTGCGGGA GGAGCGCAAG GCCTCGGACG GGTCCGTGCC GGGCTTCGGT      120
CAGCTGCACC CGCGCACGCA G                                                  141
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..138
- (D) OTHER INFORMATION: /note= "encodes porphyric herbicide resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
GCTGCAAATG CACTCTCAAA ACTATATTAC CCACCAGTTG CAGCAGTATC TATCTCGTAC      60
CCGAAAGAAG CAATCCGAAC AGAATGTTTG ATAGATGGTG AACTAAAGGG TTTTGGGCAA      120
TTGCATCCAC GCACGCAA                                                  138
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

0933123 08189

68

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: *Zea mays*
 - (B) STRAIN: B73 inbred
- (ix) FEATURE:
 - (A) NAME/KEY: -
 - (B) LOCATION: 1..138
 - (D) OTHER INFORMATION: /note= "encodes porphyric herbicide resistance domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCTGCAGATG CTCTATCAAG ATTCTATTAT CCACCGGTTG CTGCTGTAAC TGTTTCGTAT 60
 CCAAAGGAAG CAATTAGAAA AGAATGCTTA ATTGATGGGG AACTCCAGGG CTTTGGCCAG 120
 TTGCATCCAC GTAGTCAA 138

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "oligonucleotide"
- (iii) HYPOTHETICAL: NO
- (ix) FEATURE:
 - (A) NAME/KEY: -
 - (B) LOCATION: 1..36
 - (C) OTHER INFORMATION: /NOTE = "oligonucleotide primer for *Arabidopsis thaliana*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTATATTACC CACCAATGGC AGCAGTATCT ATCTCG 36

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 38 nucleotides
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: "oligonucleotide"

658180" 227EE50

69

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..38

(C) OTHER INFORMATION:/NOTE = "oligonucleotide primer for *Zea mays*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GATTCTATTA TCCACCGATG GCTGCTGTAA CTGTTTCG

38

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 nucleotides

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: "oligonucleotide"

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..26

(D) OTHER INFORMATION: /note= "oligonucleotide primer common to both of *A. thaliana* and *Z. mays* porphyric herbicide resistance domain of PPO."

(I) in addition to G, A, T or C. K = G or T, Y = C or T, S = C or G, W = A or T

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

KAYTAYCCNC CNATGGSNGC NGTNWS

26

(2) INFORMATION FOR SEQ ID NO:10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2573 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: not relevant

(ii) MOLECULAR TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Chlamydomonas reinhardtii*

(B) STRAIN: RS-3

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..2573

663T80" 22ZEE60

(C) OTHER INFORMATION: /note="encodes protoporphyrinogen oxidase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTCGAGAGCG	TTGGAGGAAA	TCCGTTTGGC	ACCTGTTCCG	GCTTCTTTGT	GTGCACGGCC	60
ACGTCCCCCT	TTCCTGCTAC	CCGCTCCCCC	CCGGCTTTAC	TGCCCCTTCC	ACTCCTCGGC	120
TCCATCCCGA	TTCCATCCGC	TCCTCCTCCC	CCACCTAGAC	TGTCTACCGT	CTACCAAGTTT	180
CTTGGGCAAT	CATTAACGTA	ACCCCGCCTC	CCTGCGCCTG	CCCCCTCCCTC	CCTCTCCCCC	240
CCGCACAGCC	CGCCGCCGCC	GAGGCCCTGG	GCTCCTTCGA	CTACCCGCCG	ATGGGCGCCG	300
TGACGCTGTC	GTACCCGCTG	AGCGCCGTGC	GGGAGGAGCG	CAAGGCCTCG	GACGGGTCCG	360
TGCCGGGCTT	CGGTCAGCTG	CACCCGCGCA	CGCAGGTGGG	CAAGTGCGCG	CGTGTTCGCG	420
GCGGTGTGTT	GCGGAGGGGA	GGGTGGTGGG	GGTTGGGGGT	GGGGGTGGGG	GGGATTGGGG	480
CGCTGGGTCTG	TATCCCGCGG	TTGTATCCTC	GCGCTCCCCT	CATCCATTCC	CCCCTTCAAC	540
AACACACACG	GGCGCACACG	CACCCTCTTT	GCGCTTACTT	TGTCTGGTGC	TCCTTAACAC	600
ACTCTTCGCT	TCATTTTGGT	GTCTTCTAAC	ACACACACTT	GTCCACACAC	AGGGCATCAC	660
CACTCTGGGC	ACCATCTACA	GCTCCAGCCT	GTTCCCCGGC	CGCGCGCCCG	AGGGCCACAT	720
GCTGCTGCTC	AACTACATCG	GCGGCACCAC	CAACCGCGGC	ATCGTCAACC	AGACCACCGA	780
GCAGCTGGTG	GAGCAGGTGT	GTGTGTGGGG	GGGTGGGGGG	GGGGCAGTGG	ATTTTGTGGG	840
TGAGCCCCCT	GAGCAAAGCG	ATCCAGGGGG	GGCGAAGCCC	CCCAGGATTG	CCCCTGTCCG	900
TGCGTGCGTG	TGTGCCTGTG	TCGACAAAAA	GTACCGTACT	GGCACAAACC	GCGAGTGCCA	960
CGTATTATTA	ATTGCAATTA	CCTATTGTAG	AAAAATAGAC	GGCAGGGAAA	ACTCGGCCGG	1020
AGCGAGAAGC	GACCTCGTGA	GTCCATGGAC	ATCTTGACTT	TCTTCAGTTC	GCGAGTATAG	1080
CTCTCGGCCC	CTAAATATCT	TACATCCATG	TATCAAAACA	TGTCGACGAC	AAGCGTCTTG	1140
GGGCAAGAAT	GTCGAAATTG	TTTGCAACAG	CCAAACCATG	CGTCCCCGAG	CCTTACATGT	1200
GTCGCGGCCC	GGGATCCCGC	GCCCGAGCCC	GGCTAGCCCT	TTGCGGTGCT	TGAGTGGGAT	1260
GTGGGTGAGG	TGCATTTGGG	ATATCATGGA	CCGTGAAGTG	GCGTGGGTAA	GGTGGCGTGG	1320
CGTGGCGGGG	ACAGGGCATG	TCGGTGCCTC	GGCACAGCGT	TGGCCTAGTG	GCCAGTCCCG	1380
CTGGATGGGC	TTGCAAGGGT	GCTGTTTCATG	TCGCCGGTGC	CCATCGTCAC	ATCCCCTTGC	1440
GCTACATGGG	GCTCAGCCCA	TTTTCCAGCT	GTACAAAGCT	GACACCCCTT	GTTGTGTGGC	1500
GTCTTGGAAC	CGTGTTGCTT	CGGAGCTGGC	CAGAACCCCC	TGTGGGCACA	CACACGCACA	1560

CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA 1620
CACACACACA CACACACACA CACACACACA CACACACACA CACATTTTCG TCCTGCAGCC 1680
CCGAACCCCG CCGCCCGTTC CACGTCTTCC ACCTGCCGCA CCCCCCCCCC TGCCGCACGC 1740
CTGCTCTCAC CGCCTCTCCC CCCACCCCAT CTCCCTGCAG GTGGACAAGG ACCTGCGCAA 1800
CATGGTCATC AAGCCCGACG CGCCCAAGCC CCGTGTGGTG GGCCTGCGCG TGTGGCCGCG 1860
CGCCATCCCG CAGGTGTGAG GGCGCAGCAG CCGGAGGGAT GGGCTAGATC CTAGTTTCTC 1920
AAAGAGCTCT ACAGCCCTAT AACCTCGACC TGCGACCTTC GACCTGATAA CCTGGCTGCC 1980
CCCTCCCAAC CTAGCCACCT CTCCCCGGAT TTGGGTTCAC TCGGTTGACT TGCTTTTGGG 2040
TTCTGGAATC AACTTCACCT GTTGATATACT TTGCTGCACT TCTCTGTACC ACTCTTTGCA 2100
TTAGGTTTCGG TTTAGTTTGG GCTGCATGTG TAACCCCTCC TCCCCGCCCT GCCACCTGCA 2160
GTTCAACCTG GGCCACCTGG AGCAGCTGGA CAAGGCGCGC AAGGCGCTGG ACGCGGCGGG 2220
GCTGCAGGGC GTGCACCTGG GGGGCAACTA CGTCAGCGGT GAGCGCGTGG GCAGCAGCAG 2280
CAGCAGGAAG AGGGGAGGGG AGGGGAGGGG AGGGTACAAG GAGGAGGTTG AGCAGGAGGT 2340
GGTGCTAAGG CGCAAAGCAA GGCGGTGTTG TATCCTCATT GACTGAAACC GGGAAACCCA 2400
GCATGAACAA GAGGTCAGGG GACTGCAAGG AGCGGAGGCT ACATGTATGA CTACCCCCGA 2460
CGCGGGCGAT GATTCCTTGA CTATTGGGAC CTATTTCGTT GGGCTCGGGC ACATGACCCC 2520
CCTGGCCCCCT TCGCTGTATG GTGCCAGCC GCCCAGCCGC CCCCCGCCCA CAC 2573

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1704 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 16..1629
- (D) OTHER INFORMATION: /product= "protoporphyrinogen

oxidase"

033173 081899

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTCTCTGCGA TTTCC ATG GAG TTA TCT CTT CTC CGT CCG ACG ACT CAA TCG	51
Met Glu Leu Ser Leu Leu Arg Pro Thr Thr Gln Ser	
1 5 10	
CTT CTT CCG TCG TTT TCG AAG CCC AAT CTC CGA TTA AAT GTT TAT AAG	99
Leu Leu Pro Ser Phe Ser Lys Pro Asn Leu Arg Leu Asn Val Tyr Lys	
15 20 25	
CCT CTT AGA CTC CGT TGT TCA GTG GCC GGT GGA CCA ACC GTC GGA TCT	147
Pro Leu Arg Leu Arg Cys Ser Val Ala Gly Gly Pro Thr Val Gly Ser	
30 35 40	
TCA AAA ATC GAA GGC GGA GGA GGC ACC ACC ATC ACG ACG GAT TGT GTG	195
Ser Lys Ile Glu Gly Gly Gly Thr Thr Ile Thr Thr Asp Cys Val	
45 50 55 60	
ATT GTC GGC GGA GGT ATT AGT GGT CTT TGC ATC GCT CAG GCG CTT GCT	243
Ile Val Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala	
65 70 75	
ACT AAG CAT CCT GAT GCT GCT CCG AAT TTA ATT GTG ACC GAG GCT AAG	291
Thr Lys His Pro Asp Ala Ala Pro Asn Leu Ile Val Thr Glu Ala Lys	
80 85 90	
GAT CGT GTT GGA GGC AAC ATT ATC ACT CGT GAA GAG AAT GGT TTT CTC	339
Asp Arg Val Gly Gly Asn Ile Ile Thr Arg Glu Glu Asn Gly Phe Leu	
95 100 105	
TGG GAA GAA GGT CCC AAT AGT TTT CAA CCG TCT GAT CCT ATG CTC ACT	387
Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Met Leu Thr	
110 115 120	
ATG GTG GTA GAT AGT GGT TTG AAG GAT GAT TTG GTG TTG GGA GAT CCT	435
Met Val Val Asp Ser Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro	
125 130 135 140	
ACT GCG CCA AGG TTT GTG TTG TGG AAT GGG AAA TTG AGG CCG GTT CCA	483
Thr Ala Pro Arg Phe Val Leu Trp Asn Gly Lys Leu Arg Pro Val Pro	
145 150 155	
TCG AAG CTA ACA GAC TTA CCG TTC TTT GAT TTG ATG AGT ATT GGT GGG	531
Ser Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Gly Gly	
160 165 170	
AAG ATT AGA GCT GGT TTT GGT GCA CTT GGC ATT CGA CCG TCA CCT CCA	579
Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg Pro Ser Pro Pro	
175 180 185	
GGT CGT GAA GAA TCT GTG GAG GAG TTT GTA CGG CGT AAC CTC GGT GAT	627
Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly Asp	
190 195 200	
GAG GTT TTT GAG CGC CTG ATT GAA CCG TTT TGT TCA GGT GTT TAT GCT	675

73

Glu 205	Val	Phe	Glu	Arg	Leu 210	Ile	Glu	Pro	Phe	Cys 215	Ser	Gly	Val	Tyr	Ala 220	
GGT Gly	GAT Asp	CCT Pro	TCA Ser	AAA Lys 225	CTG Leu	AGC Ser	ATG Met	AAA Lys	GCA Ala 230	GCG Ala	TTT Phe	GGG Gly	AAG Lys	GTT Val 235	TGG Trp	723
AAA Lys	CTA Leu	GAG Glu	CAA Gln 240	AAT Asn	GGT Gly	GGA Gly	AGC Ser	ATA Ile 245	ATA Ile	GGT Gly	GGT Gly	ACT Thr	TTT Phe 250	AAG Lys	GCA Ala	771
ATT Ile	CAG Gln	GAG Glu 255	AGG Arg	AAA Lys	AAC Asn	GCT Ala	CCC Pro 260	AAG Lys	GCA Ala	GAA Glu	CGA Arg	GAC Asp 265	CCG Pro	CGC Arg	CTG Leu	819
CCA Pro	AAA Lys 270	CCA Pro	CAG Gln	GGC Gly	CAA Gln	ACA Thr 275	GTT Val	GGT Gly	TCT Ser	TTC Phe	AGG Arg 280	AAG Lys	GGA Gly	CTT Leu	CGA Arg	867
ATG Met 285	TTG Leu	CCA Pro	GAA Glu	GCA Ala	ATA Ile 290	TCT Ser	GCA Ala	AGA Arg	TTA Leu	GGT Gly 295	AGC Ser	AAA Lys	GTT Val	AAG Lys	TTG Leu 300	915
TCT Ser	TGG Trp	AAG Lys	CTC Leu	TCA Ser 305	GGT Gly	ATC Ile	ACT Thr	AAG Lys	CTG Leu 310	GAG Glu	AGC Ser	GGA Gly	GGA Gly	TAC Tyr 315	AAC Asn	963
TTA Leu	ACA Thr	TAT Tyr	GAG Glu 320	ACT Thr	CCA Pro	GAT Asp	GGT Gly	TTA Leu 325	GTT Val	TCC Ser	GTG Val	CAG Gln	AGC Ser 330	AAA Lys	AGT Ser	1011
GTT Val	GTA Val	ATG Met 335	ACG Thr	GTG Val	CCA Pro	TCT Ser	CAT His 340	GTT Val	GCA Ala	AGT Ser	GGT Gly	CTC Leu 345	TTG Leu	CGC Arg	CCT Pro	1059
CTT Leu	TCT Ser	GAA Glu 350	TCT Ser	GCT Ala	GCA Ala	AAT Asn 355	GCA Ala	CTC Leu	TCA Ser	AAA Lys	CTA Leu 360	TAT Tyr	TAC Tyr	CCA Pro	CCA Pro	1107
GTT Val 365	GCA Ala	GCA Ala	GTA Val	TCT Ser	ATC Ile 370	TCG Ser	TAC Tyr	CCG Pro	AAA Lys	GAA Glu 375	GCA Ala	ATC Ile	CGA Arg	ACA Thr	GAA Glu 380	1155
TGT Cys	TTG Leu	ATA Ile	GAT Asp	GGT Gly 385	GAA Glu	CTA Leu	AAG Lys	GGT Gly	TTT Phe 390	GGG Gly	CAA Gln	TTG Leu	CAT His	CCA Pro 395	CGC Arg	1203
ACG Thr	CAA Gln	GGA Gly	GTT Val 400	GAA Glu	ACA Thr	TTA Leu	GGA Gly	ACT Thr 405	ATC Ile	TAC Tyr	AGC Ser	TCC Ser	TCA Ser 410	CTC Leu	TTT Phe	1251
CCA Pro	AAT Asn	CGC Arg 415	GCA Ala	CCG Pro	CCC Pro	GGA Gly	AGA Arg 420	ATT Ile	TTG Leu	CTG Leu	TTG Leu	AAC Asn 425	TAC Tyr	ATT Ile	GGC Gly	1299
GGG	TCT	ACA	AAC	ACC	GGA	ATT	CTG	TCC	AAG	TCT	GAA	GGT	GAG	TTA	GTG	1347

0933423-08189

74

Gly Ser Thr Asn Thr Gly Ile Leu Ser Lys Ser Glu Gly Glu Leu Val
 430 435 440

GAA GCA GTT GAC AGA GAT TTG AGG AAA ATG CTA ATT AAG CCT AAT TCG 1395
 Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile Lys Pro Asn Ser
 445 450 455 460

ACC GAT CCA CTT AAA TTA GGA GTT AGG GTA TGG CCT CAA GCC ATT CCT 1443
 Thr Asp Pro Leu Lys Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro
 465 470 475

CAG TTT CTA GTT GGT CAC TTT GAT ATC CTT GAC ACG GCT AAA TCA TCT 1491
 Gln Phe Leu Val Gly His Phe Asp Ile Leu Asp Thr Ala Lys Ser Ser
 480 485 490

CTA ACG TCT TCG GGC TAC GAA GGG CTA TTT TTG GGT GGC AAT TAC GTC 1539
 Leu Thr Ser Ser Gly Tyr Glu Gly Leu Phe Leu Gly Gly Asn Tyr Val
 495 500 505

GCT GGT GTA GCC TTA GGC CGG TGT GTA GAA GGC GCA TAT GAA ACC GCG 1587
 Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Thr Ala
 510 515 520

ATT GAG GTC AAC AAC TTC ATG TCA CGG TAC GCT TAC AAG TAA 1629
 Ile Glu Val Asn Asn Phe Met Ser Arg Tyr Ala Tyr Lys *
 525 530 535

ATGTAAACA TTAATCTCC CAGCTTGCCT GAGTTTATT AAATATTTTG AGATATCCAA 1689

AAAAAAAAA AAAAA 1704

(2) INFORMATION FOR SEQ ID NO:12

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 537 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN: ecotype Columbia

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..537
- (C) OTHER INFORMATION: /product="protoporphyrinogen oxidase"

(xi) SEQUENCE DESCRIPTION: SEQ. ID. NO:12:

Met Glu Leu Ser Leu Leu Arg Pro Thr Thr Gln Ser Leu Leu Pro Ser
 1 5 10 15

668780" E27EE66

75

Phe Ser Lys Pro Asn Leu Arg Leu Asn Val Tyr Lys Pro Leu Arg Leu
 20 25 30
 Arg Cys Ser Val Ala Gly Gly Pro Thr Val Gly Ser Ser Lys Ile Glu
 35 40 45
 Gly Gly Gly Gly Thr Thr Ile Thr Thr Asp Cys Val Ile Val Gly Gly
 50 55 60
 Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala Thr Lys His Pro
 65 70 75 80
 Asp Ala Ala Pro Asn Leu Ile Val Thr Glu Ala Lys Asp Arg Val Gly
 85 90 95
 Gly Asn Ile Ile Thr Arg Glu Glu Asn Gly Phe Leu Trp Glu Glu Gly
 100 105
 Pro Asn Ser Phe Gln Pro Ser Asp Pro Met Leu Thr Met Val Val Asp
 115 120 125
 Ser Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Thr Ala Pro Arg
 130 135 140
 Phe Val Leu Trp Asn Gly Lys Leu Arg Pro Val Pro Ser Lys Leu Thr
 145 150 155 160
 Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Gly Gly Lys Ile Arg Ala
 165 170 175
 Gly Phe Gly Ala Leu Gly Ile Arg Pro Ser Pro Pro Gly Arg Glu Glu
 180 185 190
 Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly Asp Glu Val Phe Glu
 195 200 205
 Arg Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser
 210 215 220
 Lys Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Lys Leu Glu Gln
 225 230 235 240
 Asn Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys Ala Ile Gln Glu Arg
 245 250 255
 Lys Asn Ala Pro Lys Ala Glu Arg Asp Pro Arg Leu Pro Lys Pro Gln
 260 265 270
 Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Arg Met Leu Pro Glu
 275 280 285
 Ala Ile Ser Ala Arg Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu
 290 295 300
 Ser Gly Ile Thr Lys Leu Glu Ser Gly Gly Tyr Asn Leu Thr Tyr Glu

0933123-081899

76

305				310				315				320			
Thr	Pro	Asp	Gly	Leu	Val	Ser	Val	Gln	Ser	Lys	Ser	Val	Val	Met	Thr
				325					330					335	
Val	Pro	Ser	His	Val	Ala	Ser	Gly	Leu	Leu	Arg	Pro	Leu	Ser	Glu	Ser
			340					345					350		
Ala	Ala	Asn	Ala	Leu	Ser	Lys	Leu	Tyr	Tyr	Pro	Pro	Val	Ala	Ala	Val
		355					360					365			
Ser	Ile	Ser	Tyr	Pro	Lys	Glu	Ala	Ile	Arg	Thr	Glu	Cys	Leu	Ile	Asp
	370					375					380				
Gly	Glu	Leu	Lys	Gly	Phe	Gly	Gln	Leu	His	Pro	Arg	Thr	Gln	Gly	Val
385					390					395					400
Glu	Thr	Leu	Gly	Thr	Ile	Tyr	Ser	Ser	Ser	Leu	Phe	Pro	Asn	Arg	Ala
				405					410					415	
Pro	Pro	Gly	Arg	Ile	Leu	Leu	Leu	Asn	Tyr	Ile	Gly	Gly	Ser	Thr	Asn
			420					425					430		
Thr	Gly	Ile	Leu	Ser	Lys	Ser	Glu	Gly	Glu	Leu	Val	Glu	Ala	Val	Asp
		435					440					445			
Arg	Asp	Leu	Arg	Lys	Met	Leu	Ile	Lys	Pro	Asn	Ser	Thr	Asp	Pro	Leu
	450					455					460				
Lys	Leu	Gly	Val	Arg	Val	Trp	Pro	Gln	Ala	Ile	Pro	Gln	Phe	Leu	Val
465					470					475					480
Gly	His	Phe	Asp	Ile	Leu	Asp	Thr	Ala	Lys	Ser	Ser	Leu	Thr	Ser	Ser
				485					490					495	
Gly	Tyr	Glu	Gly	Leu	Phe	Leu	Gly	Gly	Asn	Tyr	Val	Ala	Gly	Val	Ala
			500					505					510		
Leu	Gly	Arg	Cys	Val	Glu	Gly	Ala	Tyr	Glu	Thr	Ala	Ile	Glu	Val	Asn
		515					520					525			
Asn	Phe	Met	Ser	Arg	Tyr	Ala	Tyr	Lys	*						
	530					535									

(2) INFORMATION FOR SEO ID NO:13

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1698 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULAR TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

[illegible]

77

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Zea mays*

(B) STRAIN: B73 inbred

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..1453

(C) OTHER INFORMATION: /product="protoporphyrinogen oxidase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

G	AAT	TCG	GCG	GAC	TGC	GTC	GTG	GTG	GGC	GGA	GGC	ATC	AGT	GGC	CTC	46
Asn	Ser	Ala	Asp	Cys	Val	Val	Val	Gly	Gly	Gly	Ile	Ser	Gly	Leu		
1				5				10						15		
TGC	ACC	GCG	CAG	GCG	CTG	GCC	ACG	CGG	CAC	GGC	GTC	GGG	GAC	GTG	CTT	94
Cys	Thr	Ala	Gln	Ala	Leu	Ala	Thr	Arg	His	Gly	Val	Gly	Asp	Val	Leu	
				20				25						30		
GTC	ACG	GAG	GCC	CGC	GCC	CGC	CCC	GGC	GGC	AAC	ATT	ACC	ACC	GTC	GAG	142
Val	Thr	Glu	Ala	Arg	Ala	Arg	Pro	Gly	Gly	Asn	Ile	Thr	Thr	Val	Glu	
			35					40					45			
CGC	CCC	GAG	GAA	GGG	TAC	CTC	TGG	GAG	GAG	GGT	CCC	AAC	AGC	TTC	CAG	190
Arg	Pro	Glu	Glu	Gly	Tyr	Leu	Trp	Glu	Glu	Gly	Pro	Asn	Ser	Phe	Gln	
		50					55					60				
—e	TCC	GAC	CCC	GTT	CTC	ACC	ATG	GCC	GTG	GAC	AGC	GGA	CTG	AAG	GAT	238
Pro	Ser	Asp	Pro	Val	Leu	Thr	Met	Ala	Val	Asp	Ser	Gly	Leu	Lys	Asp	
	65					70					75					
GAC	TTG	GTT	TTT	GGG	GAC	CCA	AAC	GCG	CCG	CGT	TTC	GTG	CTG	TGG	GAG	286
Asp	Leu	Val	Phe	Gly	Asp	Pro	Asn	Ala	Pro	Arg	Phe	Val	Leu	Trp	Glu	
80					85					90					95	
GGG	AAG	CTG	AGG	CCC	GTG	CCA	TCC	AAG	CCC	GCC	GAC	CTC	CCG	TTC	TTC	334
Gly	Lys	Leu	Arg	Pro	Val	Pro	Ser	Lys	Pro	Ala	Asp	Leu	Pro	Phe	Phe	
				100					105					110		
GAT	CTC	ATG	AGC	ATC	CCA	GGG	AAG	CTC	AGG	GCC	GGT	CTA	GGC	GCG	CTT	382
Asp	Leu	Met	Ser	Ile	Pro	Gly	Lys	Leu	Arg	Ala	Gly	Leu	Gly	Ala	Leu	
			115					120					125			
GGC	ATC	CGC	CCG	CCT	CCT	CCA	GGC	CGC	GAA	GAG	TCA	GTG	GAG	GAG	TTC	430
Gly	Ile	Arg	Pro	Pro	Pro	Pro	Gly	Arg	Glu	Glu	Ser	Val	Glu	Glu	Phe	
		130					135					140				
GTG	CGC	CGC	AAC	CTC	GGT	GCT	GAG	GTC	TTT	GAG	CGC	CTC	ATT	GAG	CCT	478
Val	Arg	Arg	Asn	Leu	Gly	Ala	Glu	Val	Phe	Glu	Arg	Leu	Ile	Glu	Pro	
	145					150				155						
TTC	TGC	TCA	GGT	GTC	TAT	GCT	GGT	GAT	CCT	TCT	AAG	CTC	AGC	ATG	AAG	526
Phe	Cys	Ser	Gly	Val	Tyr	Ala	Gly	Asp	Pro	Ser	Lys	Leu	Ser	Met	Lys	
160				165				170						175		
GCT	GCA	TTT	GGG	AAG	GTT	TGG	CGG	TTG	GAA	GAA	ACT	GGA	GGT	AGT	ATT	574

668780-227EE60

78

Ala	Ala	Phe	Gly	Lys	Val	Trp	Arg	Leu	Glu	Glu	Thr	Gly	Gly	Ser	Ile	
				180					185					190		
ATT	GGT	GGA	ACC	ATC	AAG	ACA	ATT	CAG	GAG	AGG	AGC	AAG	AAT	CCA	AAA	622
Ile	Gly	Gly	Thr	Ile	Lys	Thr	Ile	Gln	Glu	Arg	Ser	Lys	Asn	Pro	Lys	
			195					200					205			
CCA	CCG	AGG	GAT	GCC	CGC	CTT	CCG	AAG	CCA	AAA	GGG	CAG	ACA	GTT	GCA	670
Pro	Pro	Arg	Asp	Ala	Arg	Leu	Pro	Lys	Pro	Lys	Gly	Gln	Thr	Val	Ala	
		210					215					220				
TCT	TTC	AGG	AAG	GGT	CTT	GCC	ATG	CTT	CCA	AAT	GCC	ATT	ACA	TCC	AGC	718
Ser	Phe	Arg	Lys	Gly	Leu	Ala	Met	Leu	Pro	Asn	Ala	Ile	Thr	Ser	Ser	
	225					230					235					
TTG	GGT	AGT	AAA	GTC	AAA	CTA	TCA	TGG	AAA	CTC	ACG	AGC	ATT	ACA	AAA	766
Leu	Gly	Ser	Lys	Val	Lys	Leu	Ser	Trp	Lys	Leu	Thr	Ser	Ile	Thr	Lys	
	240				245					250					255	
TCA	GAT	GAC	AAG	GGA	TAT	GTT	TTG	GAG	TAT	GAA	ACG	CCA	GAA	GGG	GTT	814
Ser	Asp	Asp	Lys	Gly	Tyr	Val	Leu	Glu	Tyr	Glu	Thr	Pro	Glu	Gly	Val	
				260					265					270		
GTT	TCG	GTG	CAG	GCT	AAA	AGT	GTT	ATC	ATG	ACT	ATT	CCA	TCA	TAT	GTT	862
Val	Ser	Val	Gln	Ala	Lys	Ser	Val	Ile	Met	Thr	Ile	Pro	Ser	Tyr	Val	
			275					280					285			
GCT	AGC	AAC	ATT	TTG	CGT	CCA	CTT	TCA	AGC	GAT	GCT	GCA	GAT	GCT	CTA	910
Ala	Ser	Asn	Ile	Leu	Arg	Pro	Leu	Ser	Ser	Asp	Ala	Ala	Asp	Ala	Leu	
		290					295					300				
TCA	AGA	TTC	TAT	TAT	CCA	CCG	GTT	GCT	GCT	GTA	ACT	GTT	TCG	TAT	CCA	958
Ser	Arg	Phe	Tyr	Tyr	Pro	Pro	Val	Ala	Ala	Val	Thr	Val	Ser	Tyr	Pro	
	305					310					315					
AAG	GAA	GCA	ATT	AGA	AAA	GAA	TGC	TTA	ATT	GAT	GGG	GAA	CTC	CAG	GGC	1006
Lys	Glu	Ala	Ile	Arg	Lys	Glu	Cys	Leu	Ile	Asp	Gly	Glu	Leu	Gln	Gly	
	320				325					330					335	
TTT	GGC	CAG	TTG	CAT	CCA	CGT	AGT	CAA	GGA	GTT	GAG	ACA	TTA	GGA	ACA	1054
Phe	Gly	Gln	Leu	His	Pro	Arg	Ser	Gln	Gly	Val	Glu	Thr	Leu	Gly	Thr	
				340					345					350		
ATA	TAC	AGT	TCC	TCA	CTC	TTT	CCA	AAT	CGT	GCT	CCT	GAC	GGT	AGG	GTG	1102
Ile	Tyr	Ser	Ser	Ser	Leu	Phe	Pro	Asn	Arg	Ala	Pro	Asp	Gly	Arg	Val	
			355					360					365			
TTA	CTT	CTA	AAC	TAC	ATA	GGA	GGT	GCT	ACA	AAC	ACA	GGA	ATT	GTT	TCC	1150
Leu	Leu	Leu	Asn	Tyr	Ile	Gly	Gly	Ala	Thr	Asn	Thr	Gly	Ile	Val	Ser	
			370				375					380				
AAG	ACT	GAA	AGT	GAG	CTG	GTC	GAA	GCA	GTT	GAC	CGT	GAC	CTC	CGA	AAA	1198
Lys	Thr	Glu	Ser	Glu	Leu	Val	Glu	Ala	Val	Asp	Arg	Asp	Leu	Arg	Lys	
	385					390					395					
ATG	CTT	ATA	AAT	TCT	ACA	GCA	GTG	GAC	CCT	TTA	GTC	CTT	GGT	GTT	CGA	1246

668T80"E2ZT8E60

79

Met Leu Ile Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg
400 405 410 415

GTT TGG CCA CAA GCC ATA CCT CAG TTC CTG GTA GGA CAT CTT GAT CTT 1294
Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu
420 425 430

CTG GAA GCC GCA AAA GCT GCC CTG GAC CGA GGT GGC TAC GAT GGG CTG 1342
Leu Glu Ala Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu
435 440 445

TTC CTA GGA GGG AAC TAT GTT GCA GGA GTT GCC CTG GGC AGA TGC GTT 1390
Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val
450 455 460

GAG GGC GCG TAT GAA AGT GCC TCG CAA ATA TCT GAC TTC TTG ACC AAG 1438
Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys
465 470 475

TAT GCC TAC AAG TGA TGAAAGAAGT GGAGCGCTAC TTGCCAATCG TTTATGTTGC 1493
Tyr Ala Tyr Lys *
480

ATAGATGAGG TGCCTCCGGG GAAAAAAAAG CTTGAATAGT ATTTTTTATT CTTATTTTGT 1553

AAATTGCATT TCTGTTCTTT TTTCTATCAG TAATTAGTTA TATTTTAGTT CTGTAGGAGA 1613

TTGTTCTGTT CACTGCCCTT CAAAAGAAAT TTTATTTTTC ATTCTTTTAT GAGAGCTGTG 1673

CTACTTAAAA AAAAAAAAAA AAAAA 1698

(2) INFORMATION FOR SEQ ID NO:14

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 483 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Zea mays*
- (B) STRAIN: B73 inbred

(ix) FEATURE:

- (A) NAME/KEY: peptide
- (B) LOCATION: 1..483
- (C) OTHER INFORMATION: /note="protoporphyrinogen oxidase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Asn Ser Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu Cys
1 5 10 15

003172 081899

80

Thr Ala Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu Val
 20 25 30
 Thr Glu Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg
 35 40 45
 Pro Glu Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro
 50 55 60
 Ser Asp Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp
 65 70 75 80
 Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly
 85 90 95
 Lys Leu Arg Pro Val Pro Ser Lys Pro Ala Asp Leu Pro Phe Phe Asp
 100 105 110
 Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly
 115 120 125
 Ile Arg Pro Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val
 130 135 140
 Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe
 145 150 155 160
 Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala
 165 170 175
 Ala Phe Gly Lys Val Trp Arg Leu Glu Glu Thr Gly Gly Ser Ile Ile
 180 185 190
 Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Ser Lys Asn Pro Lys Pro
 195 200 205
 Pro Arg Asp Ala Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Ala Ser
 210 215 220
 Phe Arg Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Thr Ser Ser Leu
 225 230 235 240
 Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ser
 245 250 255
 Asp Asp Lys Gly Tyr Val Leu Glu Tyr Glu Thr Pro Glu Gly Val Val
 260 265 270
 Ser Val Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala
 275 280 285
 Ser Asn Ile Leu Arg Pro Leu Ser Ser Asp Ala Ala Asp Ala Leu Ser
 290 295 300
 Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys
 305 310 315 320

09331723-081899

81

Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe
 325 330 335
 Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile
 340 345 350
 Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp Gly Arg Val Leu
 355 360 365
 Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser Lys
 370 375 380
 Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met
 385 390 395 400
 Leu Ile Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg Val
 405 410 415
 Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu
 420 425 430
 Glu Ala Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu Phe
 435 440 445
 Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu
 450 455 460
 Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys Tyr
 465 470 475 480
 Ala Tyr Lys

(2) INFORMATION FOR SEQ ID NO:15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..18

(C) OTHER INFORMATION: /note="oligonucleotide primer 1A for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15

CCGTCTACCA GTTCTTG

668180"22472260

(2) INFORMATION FOR SEQ ID NO:16

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..19

(C) OTHER INFORMATION: /note="oligonucleotide primer 1B for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16

ATACAACCGC GGGATACGA

(2) INFORMATION FOR SEQ ID NO:17

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..17

(C) OTHER INFORMATION: /note="oligonucleotide primer 2A for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17

ACTTTGTCTG GTGCTCC

(2) INFORMATION FOR SEQ ID NO:18

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

09331723-081899

83

- (ii) MOLECULAR TYPE: oligonucleotide
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: YES
- (ix) FEATURE:
 (A) NAME/KEY: -
 (B) LOCATION: 1..17
 (C) OTHER INFORMATION: /note="oligonucleotide primer 2B for *Chlamydomonas reinhardtii*"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18

TGGATCGCTT TGCTCAG

(2) INFORMATION FOR SEQ ID NO:19

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3381 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: not relevant
 (D) TOPOLOGY: not relevant
- (ii) MOLECULAR TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Chlamydomonas reinhardtii*
 (B) STRAIN: RS-3
- (ix) FEATURE:
 (A) NAME/KEY: -
 (B) LOCATION: 1..3381
 (C) OTHER INFORMATION: /note="encodes protoporphyrinogen oxidase"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CTCGAGAGCG TTGGAGGAAA TCCGTTTGGC ACCTGTTCCG GCTTCTTTGT GTGCACGGCC	60
ACGTCCCCCT TTCCTGCTAC CCGCTCCCCC CCGGCTTTAC TGCCCCCTTCC ACTCCTCGGC	120
TCCATCCCGA TTCCATCCGC TCCTCCTCCC CCACCTAGAC TGTCTACCGT CTACCAGTTT	180
CTTGGGCAAT CATTAACGTA ACCCCGCCTC CCTGCGCCTG CCCCTCCCTC CCTCTCCCCC	240
CCGCACAGCC CGCCGCCGCC GAGGCCCTGG GCTCCTTCGA CTACCCGCCG ATGGGCGCCG	300
TGACGCTGTC GTACCCGCTG AGCGCCGTGC GGGAGGAGCG CAAGGCCTCG GACGGGTCCG	360
TGCCGGGCTT CGGTCAGCTG CACCCGCGCA CGCAGGTGGG CAAGTGCGCG CGTGTTGCGG	420
GCGGTGTGTT GCGGAGGGGA GGGTGGTGGG GGTGGGGGT GGGGGTGGGG GGGATTGGGG	480

09331723-081899

CGCTGGGTCG	TATCCCGCGG	TTGTATCCTC	GCGCTCCCTT	CATCCATTCC	CCCCTTCAAC	540
AACACACACG	GGCGCACACG	CACCCTCTTT	GCGCTTACTT	TGTCTGGTGC	TCCTTAACAC	600
ACTCTTCGCT	TCATTTTGGT	GTCTTCTAAC	ACACACACTT	GTCCACACAC	AGGGCATCAC	660
CACTCTGGGC	ACCATCTACA	GCTCCAGCCT	GTTCCCCGGC	CGCGCGCCCG	AGGGCCACAT	720
GCTGCTGCTC	AACTACATCG	GCGGCACCAC	CAACCGCGGC	ATCGTCAACC	AGACCACCGA	780
GCAGCTGGTG	GAGCAGGTGT	GTGTGTGGGG	GGGTGGGGGG	GGGGCAGTGG	ATTTTTGGGC	840
TGAGCCCCCT	GAGCAAAGCG	ATCCAGGGGG	GGCGAAGCCC	CCCAGGATTG	CCCCTGTCCG	900
TGCGTGCGTG	TGTGCCTGTG	TCGACAAAAA	GTACCGTACT	GGCACAAACC	GCGAGTGCCA	960
CGTATTATTA	ATTGCAATTA	CCTATTGTAG	AAAAATAGAC	GGCAGGGAAA	ACTCGGCCGG	1020
AGCGAGAAGC	GACCTCGTGA	GTCCATGGAC	ATCTTGACTT	TCTTCAGTTC	GCGAGTATAG	1080
CTCTCGGCCC	CTAAATATCT	TACATCCATG	TATCAAAACA	TGTCGACGAC	AAGCGTCTTG	1140
GGGCAAGAAT	GTCGAAATTG	TTTGCAACAG	CCAAACCATG	CGTCCCCGAG	CCTTACATGT	1200
GTCGCGGCCC	GGGATCCCGC	GCCCCGAGCC	GGCTAGCCCT	TTGCGGTGCT	TGAGTGGGAT	1260
GTGGGTGAGG	TGCATTTGGG	ATATCATGGA	CCGTGAAGTG	GCGTGGGTAA	GGTGGCGTGG	1320
CGTGGCGGGG	ACAGGGCATG	TCGGTGCCTC	GGCACAGCGT	TGGCCTAGTG	GCCAGTCCCC	1380
CTGGATGGGC	TTGCAAGGGT	GCTGTTCATG	TCGCCGGTGC	CCATCGTCAC	ATCCCCTTGC	1440
GCTACATGGG	GCTCAGCCCA	TTTTCCAGCT	GTACAAAGCT	GACACCCCTT	GTTGTGTGGC	1500
GTCTTGGACC	CGTGTTGCTT	CGGAGCTGGC	CAGAACCCCC	TGTGGGCACA	CACACGCACA	1560
CACACACACA	CACACACACA	CACACACACA	CACACACACA	CACACACACA	CACACACACA	1620
CACACACACA	CACACACACA	CACACACACA	CACACACACA	CACATTTTCG	TCCTGCAGCC	1680
CCGAACCCCG	CCGCCCCGTT	CACGTCTTCC	ACCTGCCGCA	CCCCCCCCC	TGCCGCACGC	1740
CTGCTCTCAC	CGCCTCTCCC	CCCACCCCAT	CTCCCTGCAG	GTGGACAAGG	ACCTGCGCAA	1800
CATGGTCATC	AAGCCCCGACG	CGCCCAAGCC	CCGTGTGGTG	GGCGTGCGCG	TGTGGCCGCG	1860
CGCCATCCCG	CAGGTGTGAG	GGCGCAGCAG	CCGGAGGGAT	GGGCTAGATC	CTAGTTTCTC	1920
AAAGAGCTCT	ACAGCCCTAT	AACCTCGACC	TGCGACCTTC	GACCTGATAA	CCTGGCTGCC	1980
CCCTCCCAAC	CTAGCCACCT	CTCCCCGGAT	TTGGGTTCAC	TCGGTTGACT	TGCTTTTGGG	2040
TTCTGGAATC	AACTTCACCT	GTTGTATACT	TTGCTGCACT	TCTCTGTACC	ACTCTTTGCA	2100
TTAGGTTTCG	TTTAGTTTGG	GCTGCATGTG	TAACCCCTCC	TCCCCGCCCT	GCCACCTGCA	2160

2220
2280
2340
2400
2460
2520
2580
2640
2700
2760
2820
2880
2940
3000
3060
3120
3180
3240
3300
3360
3381

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

iii) HYPOTHETICAL: NO

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..18

(C) OTHER INFORMATION: /note="oligonucleotide primer 3A for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20

TTCCACGTCT TCCACCTG

(2) INFORMATION FOR SEQ ID NO:21

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 nucleotides

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..17

(C) OTHER INFORMATION: /note="oligonucleotide primer 3B for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21

CTAGGATCTA GCCCATC

(2) INFORMATION FOR SEQ ID NO:22

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 nucleotides

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..18

(C) OTHER INFORMATION: /note="oligonucleotide primer 4A for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22

668T80" E2ZTE60

87

CTGCATGTGT AACCCCTC

(2) INFORMATION FOR SEQ ID NO:23

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..18
- (C) OTHER INFORMATION: /note="oligonucleotide primer 4B for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23

GACCTCTTGT TCATGCTG

(2) INFORMATION FOR SEQ ID NO:24

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 nucleotides
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULAR TYPE: oligonucleotide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..17
- (C) OTHER INFORMATION: /note="oligonucleotide primer 5B for *Chlamydomonas reinhardtii*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24

CGGCATTTAC CAGCTAC

668780" E24T660